

BESZÁMOLÓK

SZEMLEK

REFERÁTUMOK

Költség és know-how – a "Máté-effektus" az információkeresésben

Az online keresések elterjedése nagymértékben hozzájárult ahhoz, hogy az információ és ára között közvetlen összefüggés létesüljön. A potenciális felhasználó azonban az információkeresést nem szívesen kapcsolja össze az információért való fizetéssel, különösképpen akkor nem, ha az ár meghatározásának módja nem világos számára. Nem szabad megfeledkezni ugyanis arról, hogy az online keresések esetében az eszközért és nem az eredményért kell fizetni, s az eszköz használatának ára, ill. az eredmény között nincs korreláció.

Az eredmény, vagyis a *visszakeresett információ* az alábbi öt kategóriába eshet:

1. az információ releváns, új és hozzáférhető;
2. az információ releváns, de a felhasználó által már ismert;
3. az információ releváns, de a felhasználó számára – intellektuális (felkészültségi vagy nyelvi) vagy fizikai okokból – nem hozzáférhető (megfelelő eszközök alkalmazásával meg lehet oldani a hozzáférés problémáját, de ha a megoldás túl nehézkes, költséges vagy bonyolult, akkor a megszerzett információhoz rendelt érték jóval alacsonyabb lehet az összes költségnél);
4. az információ érdektelen és nem releváns;
5. nincs információ.

Az első esetben a felhasználó hajlandó a szolgáltatás árának megfizetésére, a második és harmadik esetben azonban némileg becsapottnak érzi magát, s a negyedikben végképp úgy véli, hogy "kirabolták". Az ötödik eset különböző reakciókat válthat ki a felhasználóból a "semmiért való fizetés" érzésétől addig az örömig, hogy adott esetben (pl. egy új találmányra vonatkozóan) nem létezik információ. Ezért az online keresésnek lehetőleg az első kategóriába kell esnie, a második kategóriát – ha erre mód van – ki kell küszöbölni, a harmadikat kellő óvatossággal kell kezelni, a negyediknek nem szabad bekövetkeznie, s mindent meg kell tenni azért, hogy az ötödik valóban azt jelentse, amit a negatív eredmény sugall.

A felhasználó számára elfogadható árak kérdése mellett meg kell vizsgálni a költségek csökkentésének lehetőségét is. A teljesítmények maximalizálása és az elfogadható költségek együttes követelményének teljesítése a *stratégiai* és a *fogalmi* megközelítés révén lehetséges.

A stratégiai megközelítés

A hatékony stratégia első lépése a *megfelelő adatbázis(ok) kiválasztása*; ez a teljesítmény/költség hányadot jelentős mértékben befolyásolja. Ilyen megközelítésben hatékony eszköz lehet az "adatbázisok adatbázisának" ("fájlok fájljának") igénybevétele, amely általában az összes adatbázisban levő kifejezést tartalmazza adott rendszerre vonatkozóan, s amely segítségével olyan releváns adatbázis is kiválasztható, amelyre egyébként még a legképzettebb felhasználó sem gondolta.

A második fontos stratégiai elem a *keresési szoftver alapos megértése*, mivel csak az egyes utasítások alapos ismerete biztosítja a szoftver legelőnyösebb tulajdonságainak kihasználását. A menüvel ellátott rendszerek természetesen segítséget nyújtanak a felhasználóknak, ez azonban megdrágítja az igénybevételt. A parancsok csoportosítása, ill. a nem megfelelő válaszok kizárása érzékelhetően csökkenti a kereséshez szükséges időt. Fontos szerep jut a használt rendszer árképzési módjának kiaknázására épülő parancsoknak (ún. parkoló fájlok, mintanyomtatás).

A harmadik elem a *felhasználó képessége*, hogy hatékonyan legyen képes megfogalmazni a keresőkérdést. Ehhez jó segédeszköz lehet a különböző "képernyők" (osztályozási sémák, kódtáblázatok, kulcsszavak, szerző- és szervezetnevek stb.) használata.

A fogalmi megközelítés

Ez a kérdés szemantikai oldalát érinti, s a stratégiai megközelítésre van szükség akkor, ha a kérdést

leíró kifejezések nem eredményezik a keresett válasz elemeit. Ekkor minden fogalom szemantikai környezetét meg kell határozni. Igaz, hogy egy adott témakör szakértője képes fejben azonnal összeállítani egy "rögtönzött tezauszot", amelynek azonban az adatbázis által használt szógyűjtemények is meg kell felelnie. A nem szakértő ugyanakkor különböző segédeszközöket vehet igénybe a probléma megoldására.

Az adatbázisok előállítói által készített *tezauszok* pótolhatják valamelyest a felhasználó hiányos ismereteit. Különösen a kémiai, vegyipari adatbázisok tezauszai hatékonyak e tekintetben, s adott vegyület struktúrájának kanonikus megjelenítése megoldotta a nomenklatura gondját a vegyészek számára számos adatbázis esetében.

A közvetett eszközök közé tartoznak a *konkordancia listák* (a legfontosabb kulcsszavak más kulcsszavakkal együttes előfordulásait rögzítő fájlok) és a *hivatkozási fájlok* (új kifejezések, új szóhasználat felderítéséhez hasznosak).

A fogalmi megközelítéssel az adott területre szakosodott végfelhasználók nagyobb eredmény reményével foghatnak a kereséshez, de az adatbázisok használatának szakembere a segédeszközök öltetgazdag alkalmazásával nagyobb hatékonyságot tud elérni.

A "Máté-effektus" avagy a rendszerszakértő

A jó minőség/ár hányadost képviselő keresés paraméterei többdimenziósak és változóak. Csak az adatbankok és adatbázisok keresésének szakértői ismerik ezeket kellő mértékben, ill. csak ezek a szakértők tudják kellőképpen hasznosítani e paramétereket. Az alkalmi felhasználó – bármennyit tud is az illető szakterületről – nem képes hasonló teljesítményre. Csak a legalább 100 adatbázisban

keresést végzett szakember képes a legígéretesebb fájlok kiválasztására, ill. az eredmények kiegészítésére a kevésbé kézenfekvő fájlokból. A keresési szakértő gyorsabban reagál, kifinomultabb stratégiát alkalmaz és még a legegyszerűbb esetben is 25%-kal gyorsabban és sokkal jobb eredményt ér el a szoftver tökéletesebb ismerete révén.

A "Máté-effektus" ("Mert annak, akinek van, még adnak, hogy bőven legyen neki; akinek meg nincs, attól még amije van is, elveszik" – olvasható Máté evangéliumában) ebben a vonatkozásban azt jelenti, hogy a magas színvonalú információkereséseket végző szakemberhez mind gyakrabban fordulnak. Következésképpen egyre tapasztaltabbá és eredményesebbé válik, ezért több keresést bíz rá, s egyre jobban végzi ezt a tevékenységet. (A "Máté-effektust" a szakirodalom eddig az "információban szegények" és az "információban gazdagok" közötti különbség leírására használta. – A ref.)

Az eseti felhasználó – aki eredményeit a ráfordított idő, a keresés eredményeként nyert válaszok és a kapott számlák alapján értékeli – rájön arra, hogy jobban jár, ha a keresést a szakemberre bízta. Mindenesetre ezt látszanak bizonyítani az *Institut Français du Pétrole* online kereséssel kapcsolatos tapasztalatai. A mennyiségét és komplexitását tekintve növekvő kereslet, az egyre komplexebbé váló fájlok és a kötött költségvetés körülményei között csak így lehetett a szolgáltatást az igényeknek megfelelően üzemeltetni az intézménynél. Ehhez azonban centralizált szervezetre van szükség a teljesítmények és a költségek figyelemmel kísérése érdekében.

/MOUREAU, M.: Cost and know-how – the "Matthew effect" in information retrieval. = Online Review, 11. köt. 6. sz. 1987. p. 355–360./

(Hegedűs Péter)

Bevezetés a numerikus adatbázisok világába

Sokan közülünk félnek a számoktól. Az információ szakembernek azonban nem szabad ebbe a csapdába esnie. A numerikus adatokkal is kényelmesen és eredményesen dolgozhatunk, anélkül, hogy a számok szakemberei lennénk. A numerikus adatbázisokat arra tervezték, hogy nagy adatmennyiségeket lehessen velük hatékonyan kezelni, használatukat pedig semmivel sem nehezebb megtanulni, mint a bibliográfiai adatbázisokét.

Mik azok a numerikus adatbázisok?

A numerikus adatbázisok a nem bibliográfiai (vagy forrás-) adatbázisok körébe tartoznak. Az ilyen adatbázisok magát a keresett információt tartalmazzák, míg a bibliográfiai adatbázisok azt, hogy az információ

hol lelhető fel*. A nem bibliográfiai típusú adatbázisok körébe tartoznak a teljes szövegű adatbázisok (pl. az *INVESTEXT*), az útmutatók (pl. az *Electronic Yellow Pages*), a grafikus adatbázisok (pl. a *Chemical Abstracts* szerkezetikeplet-állománya), valamint tárgyunk, a numerikus adatbázisok.

A tisztán numerikus adatbázisok, amelyekkel foglalkozni akarunk, a számadatok mellett csak nagyon kevés – meghatározás jellegű – szöveges információt tartalmaznak. Ezek az adatbázisok elsősorban a

* Az elhatárolás nem mindig ilyen egyszerű. Egyes adatbázisok, mint például a *Predicasts PROMT* és a *Metals Data File* együtt tartalmazzák magukat a numerikus adatokat és forrásuk bibliográfiai adatait. – A ref.